Protein Engineering, Design & Selection vol. 25 no. 5 pp. 205–212, 2012
Published online February 20, 2012 doi:10.1093/protein/gzs007

The intracellular Ig fold: a robust protein scaffold for the engineering of molecular recognition

Marc Bruning[1,2], Igor Barsukov[1], Barbara Franke[1], Sonia Barbieri[1,2], Martin Volk[3], Sonja Leopoldseder[2], Zöhre Ucurum[2] and Olga Mayans[1,4]

[1]Institute of Integrative Biology, University of Liverpool, Crown Street, Liverpool L69 7ZB, UK, [2]Division of Structural Biology, Biozentrum, University of Basel, Klingelbergstr. 70, Basel CH-4056, Switzerland and [3]Surface Science Research Centre, Department of Chemistry, University of Liverpool, Liverpool L69 7ZD, UK

[4]To whom correspondence should be addressed.
E-mail: Olga.Mayans@liv.ac.uk

Received August 30, 2011; revised January 10, 2012; accepted January 17, 2012

Edited by Jane Clarke

Protein scaffolds that support molecular recognition have multiple applications in biotechnology. Thus, protein frames with robust structural cores but adaptable surface loops are in continued demand. Recently, notable progress has been made in the characterization of Ig domains of intracellular origin—in particular, modular components of the titin myofilament. These Ig belong to the I(intermediate)-type, are remarkably stable, highly soluble and undemanding to produce in the cytoplasm of *Escherichia coli*. Using the Z1 domain from titin as representative, we show that the I-Ig fold tolerates the drastic diversification of its CD loop, constituting an effective peptide display system. We examine the stability of CD-loop-grafted Z1-peptide chimeras using differential scanning fluorimetry, Fourier transform infrared spectroscopy and nuclear magnetic resonance and demonstrate that the introduction of bioreactive affinity binders in this position does not compromise the structural integrity of the domain. Further, the binding efficiency of the exogenous peptide sequences in Z1 is analyzed using pull-down assays and isothermal titration calorimetry. We show that an internally grafted, affinity FLAG tag is functional within the context of the fold, interacting with the anti-FLAG M2 antibody in solution and in affinity gel. Together, these data reveal the potential of the intracellular Ig scaffold for targeted functionalization.

Keywords: differential scanning fluorimetry/infrared spectroscopy/intracellular Ig domain/NMR spectroscopy/protein scaffold engineering for peptide display

Introduction

Protein scaffolds with adaptable molecular recognition properties are widely used in biotechnology and biomedicine, e.g. in medical imaging, immunoassays, microarrays and intracellular therapy. Many applications rely on antibodies, antibody fragments or their engineered variants (e.g. Fab, single-chain variable fragments (scFv), diabodies, triabodies and minibodies) that have been used as archetypal binders for decades (Holliger and Hudson, 2005; Lo *et al*., 2008; Wurch *et al*., 2008; Nelson and Reichert, 2009). Antibodies are naturally produced by the immune system through a process of hypermutation that yields myriads of variants with unique binding attributes. This versatility relies on the structural plasticity of the Ig fold of antibody domain components. The Ig fold (~10 kDa) consists of two amphipathic β-sheets that pack against each other to form a β-sandwich. Several fold variants are found in proteins that differ in β-strand composition and loop architecture, forming the Ig-like superfamily (Halaby *et al*., 1999). Among these, the V(variable)-Ig type mediates antigen recognition in antibodies. This Ig has three hypervariable loops (BC, C′C″, FG) at the N-terminal pole of the β-sandwich, which can accommodate large differences in sequence, length and conformation, being efficient display loci for binder motifs. Thus, the V-Ig fold is regarded as the canonical protein scaffold.

Antibodies can be developed through immunization or *in vitro* display methods to exhibit high binding affinity and specificity. However, their complex architecture (requiring glycosylation and disulfide bond formation) makes them costly to manufacture, needing stably transfected mammalian cell lines (Steinmeyer and McCormick, 2008). Antibody fragments and their engineered variants, that have simplified structures and can be optimized through design, are more amenable to economical bacterial expression. These smaller binders also have functional advantages, being able to access cryptic epitopes and allowing the production of immunoconjugates by fusion to molecules such as enzymes for pro-drug therapy, toxins for cancer treatment and liposomes for drug delivery (Holliger and Hudson, 2005; Nelson and Reichert, 2009). Fragments, however, often suffer aggregation due to the hydrophobicity of their surface. This problem can be alleviated by surface mutagenesis but success is variable (Nelson and Reichert, 2009). Interest in small Ig binders was boosted by the discovery that camelids and cartilaginous fish have high-affinity single V-Ig that can be produced as soluble and stable *in vitro* products, and whose humanization makes them promising for clinical applications (Nelson and Reichert, 2009).

Difficulties with the production and intellectual property context of Ig-based binders triggered a search for alternative protein scaffolds. Non-Ig frames are now available that can be re-engineered through combinatorial protein design and *in vitro* selection to exhibit tailored binding functions (Wurch *et al*., 2008; Gebauer and Skerra, 2009; Grönwall and Ståhl, 2009; Löfblom *et al*., 2011). Examples of new scaffolds are: affibodies derived from protein A, lipocalin-based anticalins, designed ankyrin repeat domains, knottins and adnectins. The latter, derived from the 10th domain of fibronectin, have a Fn3 fold that also belongs to the Ig-like superfamily, being the only member with scaffold properties other than the V-Ig. The advantages of these new frames are (i) their low manufacturing cost; (ii) the ease of generating multispecific

molecules; and (iii) their strong activity in *in vitro* models. These proteins replace and even surpass antibody performance in applications such as medical imaging and diagnostic protein microarrays, and several have progressed to clinical testing showing promise as biotherapeutics (Wurch *et al*., 2008; Grönwall and Ståhl, 2009; Löfblom *et al*., 2011). Thus, the relevance of alternative scaffolds is expected to continue increasing.

Intracellular variants of the Ig fold that belong to the I(intermediate)-type (Halaby *et al*., 1999) are found as modular components of muscle filaments and cytoskeletal proteins, where they occur in high number (Kontrogianni-Konstantopoulos *et al*., 2009). A prominent example of an intracellular poly-Ig protein is the titin myofilament that contains >300 Ig and Fn3 domains linked in tandem (Labeit and Kolmerer, 1995). Titin domains offer ease of production in bacterial systems, and are highly soluble and highly stable. In recent years, significant knowledge of their atomic structure has been gained, with 3D models now available for 17 Ig and 4 Fn3 modules from titin. Namely (from N- to C-terminus of titin), Ig domains Z1 and Z2 (Marino *et al*., 2006; Zou *et al*., 2006), I1 (Mayans *et al*., 2001), I65–I70 (von Castelmur *et al*., 2008), I91 (previously I27) (Improta *et al*., 1996), A164 and A165 (PDB entry 3LCY), A168 and A169 (Mrosek *et al*., 2007; Müller *et al*., 2007), M1 (2BK8), M5 (Pfuhl and Pastore, 1995) and M10 (Pernigo *et al*., 2010); and Fn3 domains A71 (Goll *et al*., 1998), A77 and A78 (Bucher *et al*., 2010) and A170 (Mrosek *et al*., 2007). These folds are not stabilized by disulfide bonds but by robust hydrophobic cores that confer them stability in the intracellular environment. Supporting the expectation that these domains might be of technological interest, Z1–Z2 have been recently employed in the engineering of self-assembling protein nanofibers (Bruning *et al*., 2010) and the tandem I65–I70 has been shown to form layers on liposomes (Chami *et al*., 2011). In contrast, obtaining higher order assemblies using antibodies requires polymeric environments: e.g. the entrapment of antibody-coated beads in hydrogels for the generation of protein microarrays (Charles *et al*., 2004). These approaches allow little control over the distribution of functions in a material. In this regard, the I-Ig type might offer the ease of introducing molecular recognition into self-organizing systems. Further, intracellular binders that track or block endogenous proteins are valuable tools for the manipulation of cellular processes. For example, redesigned scFv intrabodies are applied in cancer, autoimmune and neurodegenerative disease and transplantation (Lo *et al*., 2008), as well as in cellular imaging (Grönwall and Ståhl, 2009). However, the development of intrabodies is troubled by the difficulty of imposing an *in vitro* selection pressure for intracellular stability (Lo *et al*., 2008). The intracellular Ig frame might be valuable in this context.

Yet, most Ig from titin play structural roles and those few known to mediate protein interactions do not utilize V-Ig equivalent loops. For example, Z1–Z2 and M10 interact with their respective protein partners, telethonine and obscurin-like-1, by forming an inter-molecular β-sheet through their β-strand G (Zou *et al*., 2006; Pernigo *et al*., 2010); and domain A169 binds MuRF1 through a unique loop in the A–A′ region (Mrosek *et al*., 2007). In fact, the N-terminal, hypervariable loops of V-Ig are short and conserved in the I-Ig type of titin (Marino *et al*., 2005), so that their modification might be poorly tolerated. Accordingly, a study that modified I91 (ex I27) by introducing five glycine residues in the BC loop showed that this did not affect the overall fold, but caused a large destabilization (Wright *et al*., 2004). Thus, it is unclear whether this Ig type can mediate interactions through its loop regions. Using the well-characterised Z1 domain as representative, we have explored the potential of the intracellular Ig type to act as an adaptable scaffold in mediating protein interactions.

Materials and methods

Sequence conservation

The 3D structures of titin domains Z1 and Z2 (PDB entry 2A38), I1 (1G1C), I65–I70 (3B43), A164 (3LCY), A168 (2NZI), M1 (2BK8) and M10 (2Y9R) were aligned using PDBeFOLD (Krissinel and Henrick, 2004) that matched their secondary structure (root mean square deviation <1.5 Å for all models). Next, the sequences of all other Ig from titin (as defined in entry Q8WZ42, UniProtKB) were aligned against the structure-based alignment using PROMALS3D (Pei *et al*., 2008). Conservation percentages were calculated in PROTSKIN (Deprez *et al*., 2005) with the BLOSUM62 matrix and mapped onto the Z1 structure using 10 color levels corresponding to 10% conservation increments.

Cloning

Cloning of titin Z1 (residues 1-99; Q8WZ42) was in the pETM-11 vector (EMBL plasmid collection) using KpnI and NcoI restriction sites. This vector incorporates a His_6-tag and a tobacco etch virus (TEV) protease cleavage site N-terminal to the target construct. Z1 variants were constructed from the wild-type clone using overlap extension PCR, where sequences were introduced in Z1 between residues D42 and L51. For this, first, two DNA fragments were constructed coding for Z1 residues 1–42 and 51–99 and containing additional codons for the inserted sequences at the respective 3′ or 5′ termini. Then, the overlap in the added sequences of the resulting DNA fragments generated the entire Z1 variant construct. All clones were confirmed by sequencing.

The expression clone of the SH3 domain from Pexp13 has been previously reported (Douangamath *et al*., 2002).

Protein production

Recombinant overexpression of Z1 and its variants was carried out in *Escherichia coli* Rosetta (Novagen) using a same protocol. Cells were grown in Luria–Bertani medium supplemented with 25 μg/ml kanamycin and 34 μg/ml chloramphenicol at 37°C up to an OD_{600} of 0.6–0.8. Protein expression was induced with 1 mM isopropyl β-D-thiogalactoside and cultures grown further at 20°C overnight for a minimum of 12 h. Harvesting of the cells was by centrifugation. For purification, cell pellets were resuspended in a lysis buffer containing 50 mM Tris-HCl pH 7.4, 100 mM NaCl and supplemented with 1 mg/ml lysozyme, 0.01 mg/ml Dnase I. Cells were lysed by sonication. Lysates were clarified by centrifugation and supernatants were applied to a Ni^{2+}-nitrilotriacetate Hi-Trap column (GE Healthcare) equilibrated in lysis buffer. Elution used 200 mM imidazole. Eluted fractions were dialyzed against lysis buffer in the presence of TEV protease and 1 mM

dithiothreitol for His_6-tag removal. Further purification was by substractive metal affinity and size-exclusion chromatography in a Superdex 75 HR16/60 column (GE Healthcare) equilibrated in lysis buffer. Purified samples were concentrated to 1 mM and stored at 4°C until further use.

Differential scanning fluorimetry

The thermal denaturation of Z1 and its variants was recorded by differential scanning fluorimetry (DSF) in triplicate on a 7500 Fast Real Time PCR System (Applied Biosystems) in the presence of Sypro Orange (Invitrogen) using a wavelength of 560 and 582 nm for fluorescence excitation and detection, respectively (Niesen *et al*., 2007). The concentration of all protein samples was 50 μM in a final sample volume of 10 μl in Tris-HCl pH 7.4, 100 mM NaCl. A Sypro Orange 100× solution was added at 10% (v/v) in a Fast Optical 96 Well Reaction Plate (Applied Biosystems). The heating cycle comprised a 120-s prewarming step at 31°C and a subsequent gradient 32–81°C obtained in 99 steps of 20 s (0.5°C ramp). Human acidic fibroblast growth factor was used as a control, showing a $T_m = 48$°C in good agreement with previously reported values (Chi *et al*., 2001). Data were corrected for background fluorescence at time $t = 0$ and normalized using GNUPLOT (www.gnuplot.info). T_m values and standard deviations were obtained from the first derivative of the triple DSF runs.

Fourier transform infrared spectroscopy

Prior to the measurements, buffers were prepared in 100% D_2O to consist of 100 mM deuterated phosphate buffer pD 7.6, 100 mM NaCl. Sample media were exchanged thoroughly against these buffers by sequential dialysis over a 16-h period. Protein concentrations were: Z1 13.6 mg/ml, $Z1^{PxxP}$ 13.7 mg/ml, $Z1^{FLAG}$ 7.8 mg/ml. Fourier transform infrared (FTIR) spectra were recorded at 10°C on a Bio-Rad FTS-40 FTIR spectrometer equipped with a liquid nitrogen-cooled HgCdTe detector using a temperature-controlled IR-cell with CaF_2 windows and 50 μm spacer. For each spectrum, 500 scans with 1 cm^{-1} resolution were averaged. All spectra were corrected to yield a flat spectrum in the region 2000–1800 cm^{-1} by subtracting an appropriately scaled buffer spectrum.

Nuclear magnetic resonance spectroscopy

Uniform labeling with ^{15}N ($Z1^{FLAG}$) or ^{15}N and ^{13}C (Z1) was carried out by recombinant expression in M9 medium supplemented with 1 g/l $^{15}NH_4Cl$ and/or 2 g/l [$^{13}C_6$]-glucose as nitrogen and carbon source, respectively. Production protocols were otherwise as those of unlabeled samples. Samples were concentrated to 1 mM in 20 mM phosphate buffer pH 6.0, 20 mM NaCl, 5% (v/v) D_2O. Nuclear magnetic resonance (NMR) measurements were carried out at 25°C on a Bruker ADVANCE DRX 600 equipped with a CryoProbe. Referencing of the proton chemical shifts was to external DSS, while the ^{15}N and ^{13}C chemical shifts were indirectly referenced using recommended gyromagnetic ratios (Wishart *et al*., 1995). All spectra were processed and analysed using the software TopSpin (Bruker) and CCPN (Vranken *et al*., 2005), respectively. For the sequential assignment of backbone NH, N, CO, Cα and Cβ resonances the following experiments were recorded for wild-type Z1 and $Z1^{FLAG}$: 3D HNCO, 3D HNCACB, 3D CACB(CO)NH. Initial values for the chemical shifts were obtained from the Z1–Z2 tandem (BioMagResBank, entry 5760; Edlich and Muhle-Goll, 2003), and adjusted for Z1. Due to the large differences between the [^1H,^{15}N]-HSQC of Z1 and $Z1^{FLAG}$, the resonances of $Z1^{FLAG}$ were assigned independently using the 3D spectra. The degree of broadening in the [^1H,^{15}N]-HSQC spectrum of $Z1^{FLAG}$ was estimated from the ratio I_{Z1FLAG}/I_{Z1} of the cross-peak intensities. The ratio was normalized to give an average value of 1 for the non-broadened resonances. The difference between the average backbone ^{13}C chemical shift values of Z1 and $Z1^{FLAG}$ was calculated as $(\Delta CO^2 + \Delta C\alpha^2 + \Delta C\beta^2)^{-1/2}/3$. The dihedral angles for both samples were predicted from the chemical shift values using the DANGLE (Cheung *et al*., 2010) module in CCPN (Vranken *et al*., 2005). Backbone ^{13}C, ^{15}N and $^1H^\alpha$ chemical shifts were used in the dihedral angle prediction.

Anti-FLAG binding assay

The binding of the FLAG peptide in $Z1^{FLAG}$ was tested by pull-down using anti-FLAG M2 affinity gel (Sigma). Wild-type Z1 was employed as negative control. Both Z1 and $Z1^{FLAG}$ were used at a concentration of 0.1 mM in 50 mM Tris-HCl pH 7.4, 100 mM NaCl. In independent experiments, each sample was applied to 200 μl of anti-FLAG resin. The resin was then washed four times with sample buffer and bound protein eluted with 1 mM FLAG peptide (Sigma). All fractions were monitored by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE).

Isothermal titration calorimetry

Samples were thoroughly dialyzed against 50 mM Tris-HCl pH 7.4, 100 mM NaCl, 0.05% NaN_3 and data then recorded with an iTC_{200} calorimeter (Microcal) at 25°C. Titrations consisted of 2 μl injections of $Z1^{FLAG}$ concentrated to 0.1 mM into anti-FLAG M2 antibody (SIGMA F1804) solutions at 0.01 mM. The time interval between injections was 200–300 s. The data were corrected for the heat of dilution by subtracting the small constant heat obtained at the end of the titration, where no further binding occurs.

Results and discussion

Fold conservation in intracellular Ig domains from titin

To reveal regions of natural variability in the I-Ig fold of titin, we studied the sequence conservation of its Ig modules (Supplementary Fig. S1) and mapped conservation levels onto the 3D structure of Z1 (Fig. 1). This exposed two main areas of fold preservation: (i) the hydrophobic core and (ii) loops BC and FG in the N-terminal apex of the fold. The hydrophobic core primarily consists of an almost strictly conserved tryptophan residue, W39 in β-strand C, and residues L65 and Y78 in β-strands E and F, respectively, which are 89 and 93% conserved. Residue W39 belongs to the signature motif of canonical Ig domains (Ioerger *et al*., 1999), while Y78 forms the classical 'tyrosine corner' (Hamill *et al*., 2000). Both residues are highly conserved across the Ig superfamily and are important for folding and stability (Halaby *et al*., 1999). The conservation in the N-terminal loop region (hypervariable in V-Ig domains) involves a Pro residue from the BC loop (P33) and an NxxG motif in the

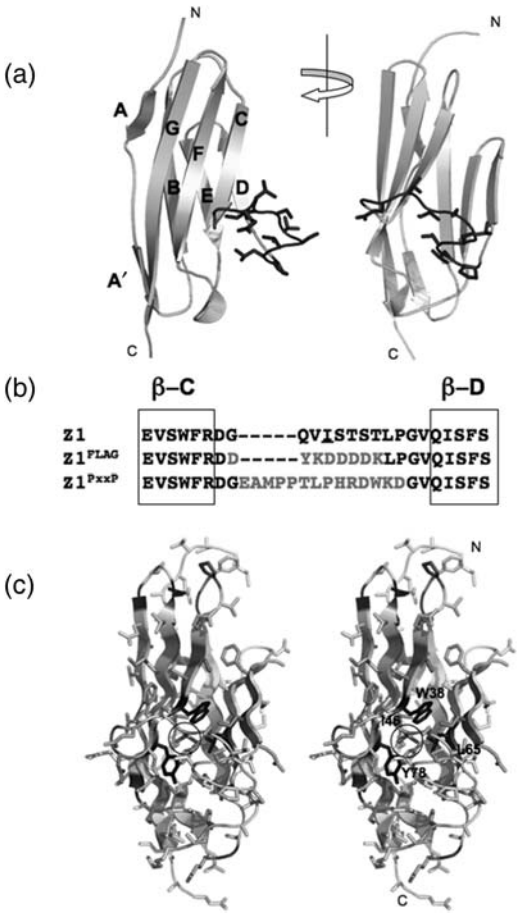

Fig. 1. The Z1 domain. (**a**) Crystal structure of Z1. Secondary structure elements are labelled. The protruding CD loop undergoing grafting is shown in dark grey; (**b**) Sequence spanning the CD loop in Z1 and the engineered variants $Z1^{FLAG}$ and $Z1^{PxxP}$ (inserts are in grey); (**c**) Sequence conservation derived from the alignment of all Ig domains from titin mapped onto the 3D structure of Z1. A white-to-black gradient is used to indicate low (<10% seq. id; white) to high (>90% seq. id; black) conservation levels. (This figure is provided in color as Supplementary Fig. S4.)

FG β-hairpin. These residues define the packing of the N-terminal loop cluster and their structural role in the I-Ig fold has been previously described (Marino *et al.*, 2005).

The most variable parts in these Ig are loops CD and EF, at the C-terminal pole of the β-sandwich, which are conserved in length but vary in sequence. While the EF loop packs against the A′B connector, the CD loop (residues 43–53 in Z1) protrudes from the β-sandwich and is highly exposed (Fig. 1a). The CD loop is orthogonal to the longitudinal axis of the fold, being also accessible within the context of the natural poly-Ig arrays of titin. Only one position in this loop is conserved, residue I46, which is always hydrophobic and in 91% of the modules is the isomeric rest Ile/Leu. This position is spatially close to the core residues W39 and Y78 (Fig. 1c), and contributes to the packing of the CD loop against the β-sandwich, suggesting that it might be important for fold stability. However, overall, the variability and accessibility of the CD loop indicates that this is a good candidate for engineering purposes.

Viable Z1 variants obtained by loop grafting

To explore the potential of the CD loop to display exogenous peptide sequences, we grafted two established recognition motifs on the corresponding region of Z1. Two variants, $Z1^{FLAG}$ and $Z1^{PxxP}$, were constructed (Fig. 1b). In $Z1^{FLAG}$, eight naturally occurring residues were replaced by the FLAG sequence (DYKDDDDK), which is a conventional affinity tag. In $Z1^{PxxP}$, 9 native residues were substituted for a 14 amino acid (aa) sequence (EAMPPTLPHRDWKD) containing a canonical PxxP motif for SH3 domain recognition. The sequence used here is a segment from the peroxisomal protein Pex14p that interacts tightly with the SH3 domain from Pex13p (Douangamath *et al.*, 2002). The selection of peptides in this study was directed to probe two contrasting scenarios regarding loop composition and conformation. In FLAG, the sequence is highly charged and it is expected to form an unstructured, flexible loop within Z1. On the contrary, the PxxP peptide has a high hydrophobic content and stringent binding requirements, where the recognition sequence spans eight aa that must form an extended arrangement of ∼22 Å length in PPII conformation (Douangamath *et al.*, 2002).

Both $Z1^{FLAG}$ and $Z1^{PxxP}$ variants were undemanding to produce in *E.coli* and yielded 30–40 mg of pure protein per 1 l culture, resembling yields obtained for wild-type Z1. The size-exclusion chromatograms of variants and wild type were also similar (Fig. 2a), indicating that the grafted variants retain the monomeric character of the domain. To estimate the effect of the modifications on fold stability, the thermal denaturation of Z1, $Z1^{FLAG}$ and $Z1^{PxxP}$ was monitored using DSF (Fig. 2b). The melting temperature (T_m) of wild-type Z1 was 72.6 ± 0.16°C, in good agreement with CD data previously reported ($T_m = 69.4 \pm 0.1$°C) (Politou *et al.*, 1995). Variants $Z1^{FLAG}$ and $Z1^{PxxP}$ yielded T_m values of 60.48 ± 0.08°C and 58.94 ± 0.74°C, respectively. While their $\Delta T_m \approx 10$°C indicated that the alteration of the CD loop was destabilizing, the overall thermal stability of the variants still resembled (and even surpassed) that of other naturally occurring Ig domains in titin (Politou *et al.*, 1995; Mayans *et al.*, 2001).

The reduction in stability led us to investigate a possible perturbation of the secondary structure of the fold. For this, we studied the FTIR spectra of wild type and variants (Fig. 2c). The amide I′ band (prime indicates fully deuterated amide groups) from all samples presented a low maximal frequency (1633 cm^{-1} for Z1 and $Z1^{PxxP}$, and 1635 cm^{-1} for $Z1^{FLAG}$), characteristic of β-rich proteins (Barth, 2007). Side-chain contributions from Gln and Asn residues (1635–1650 cm^{-1}), Arg (1608 and 1586 cm^{-1}) or Tyr and Trp (∼1615 cm^{-1}) (Barth, 2007) should be negligible due to the low (and nearly identical) content of these amino acids in the samples (Gln/Asn ∼7%; Arg 3–4%; Tyr/Trp < 3%). Thus, the FTIR spectra could be comparatively interpreted in terms of secondary structure content. The spectra of Z1 and $Z1^{PxxP}$ were virtually identical, indicating a highly similar secondary structure. The $Z1^{FLAG}$ spectrum showed a small difference compared with that of Z1 and $Z1^{PxxP}$ in that the maximum of its amide I' band was shifted towards a slightly higher value. An influence from changes in side-chain absorbance can be excluded; as these would have the opposite effect (the replacement of Gln by Tyr would shift the band to lower frequencies). The shift to higher values is likely to derive from a slight loosening of the β-strands in $Z1^{FLAG}$ compared with Z1 and $Z1^{PxxP}$. Still, the shift is very small and the secondary structure of $Z1^{FLAG}$ can be deduced to be closely similar to that of Z1.

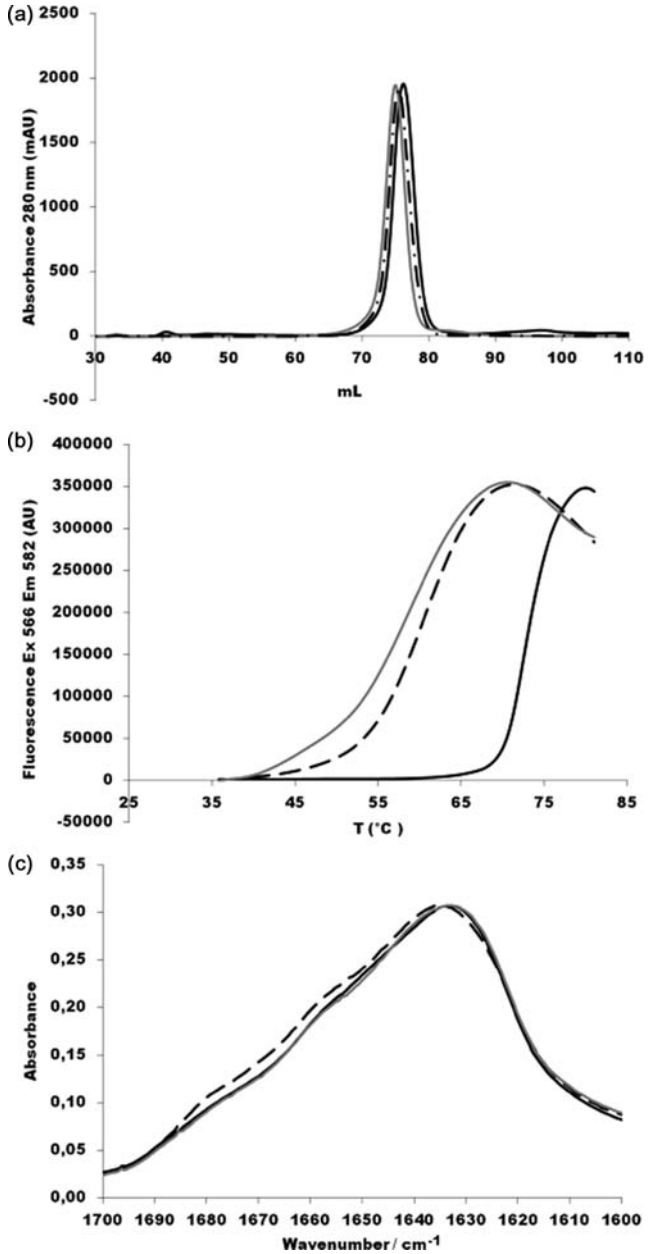

Fig. 2. Z1 fold stability. (**a**) Size-exclusion chromatograms in 50 mM Tris-HCl pH 7.4, 100 mM NaCl at sample concentrations of ~5 mg/ml. Exclusion volumes were: 76.14 ml for Z1, 74.97 ml for $Z1^{FLAG}$ and 75.45 ml for $Z1^{PxxP}$; (**b**) DSF curves showing the increase in the unfolded protein fraction as a function of temperature; (**c**) FTIR spectra of the amide I' region (the spectra of $Z1^{PxxP}$ and $Z1^{FLAG}$ were scaled to that of Z1 to yield a same maximum absorbance). Throughout, a solid black line indicates Z1, dashed line is $Z1^{FLAG}$ and solid grey line is $Z1^{PxxP}$.

Taken together, data from size-exclusion chromatography, DSF and FTIR suggest that the modification of the CD loop has been well tolerated by Z1 and that the contribution of I46 to fold stability is dispensable. The diverse nature of the grafted sequences shows that the effect is general and that the loop can accommodate contrasting amino acid compositions as well as long insertions beyond the native length of the loop.

Localization of structural alterations in $Z1^{FLAG}$ by NMR

The small structural changes in $Z1^{FLAG}$ relative to Z1 and $Z1^{PxxP}$ prompted a detailed analysis of this variant by NMR. As reference, the [^{1}H,^{15}N]-HSQC spectrum of Z1 was recorded (Fig. 3a and Supplementary Fig. S2). This contained well-dispersed peaks of uniform line width as expected for a small rigid globular protein. The chemical shifts of the isolated Z1 were similar to those previously reported for Z1 within the Ig–Ig pair Z1–Z2 (Edlich and Muhle-Goll, 2003). The [^{1}H,^{15}N]-HSQC spectrum of $Z1^{FLAG}$ (Fig. 3b and Supplementary Fig. S2) showed an overall chemical shift dispersion similar to that of Z1, indicating that it has an equivalent globular fold. However, a few peaks in this spectrum had broadened noticeably. Comparative spectra recorded at significantly lower protein concentration (0.01 mM instead of 1 mM otherwise employed in this work) and increased ionic strength (100 mM NaCl instead of 20 mM used otherwise) demonstrated no change in resonance broadening or significant variation in chemical shifts (Supplementary Fig. S3). This suggested that resonance broadening is caused by an internal dynamic process localized to a small region of the fold, and not by intermolecular association. In the latter, a concentration dependence of chemical shifts of the broadened resonances would be expected, as well as a reduction of the broadening at low concentrations.

Most of the $Z1^{FLAG}$ signals could be assigned using triple-resonance experiments, except for sequences E36-F40 and V54-S58 (β-strands C and D, respectively) that flank the modified loop (Fig. 3c). We concluded that the missing signals corresponded to the severely broadened resonances that remained unassigned. This identified β-strands C and D as the main regions undergoing conformational exchange that is intermediate on the NMR time scale. This suggested a loosening of those strands, in agreement with FTIR data. The strands must remain associated to their respective β-sheets as their detachment would have resulted in a sharpening (not a broadening) of their signals. Their detachment would have also caused more significant changes in the FTIR-amide I' band. The loosening of β-strands C and D in $Z1^{FLAG}$ might be primarily caused by the removal of the hydrophobic group I46, which is wedged between the strands and possibly stabilizes their mutual packing (Fig. 1c).

Next, the degree of broadening of the assigned resonances in the $Z1^{FLAG}$ spectrum was estimated from the ratio I_{Z1FLAG}/I_{Z1} calculated for each cross peak. This ratio was normalized to give an average value of 1 for non-broadened resonances. An increase in I_{Z1FLAG}/I_{Z1} ratio indicated a narrowing of the line-width of a given $Z1^{FLAG}$ signal relative to that of Z1, while a decrease signified resonance broadening. The ratio was close to 1 for the majority of the assigned cross peaks, suggesting that the alteration of the CD loop had little effect on most residues. Only a small number of resonances showed significant broadening ($I_{Z1FLAG}/I_{Z1} < 0.4$). Mapping the corresponding residues onto the Z1 structure (Fig. 3c) revealed that these were located next to β-strands C and D, or connected to them through a network of contacts, indicating that the perturbation of the scaffold by the loop modification was restricted to the immediate vicinity of the latter. In contrast, the introduced sequence showed increased mobility ($I_{Z1FLAG}/I_{Z1} > 1$). In Z1, the intensities of the CD loop resonances were similar to those in the β-strands, suggesting that the loop is immobilized and in close contact with the rest of the structure (as confirmed by the

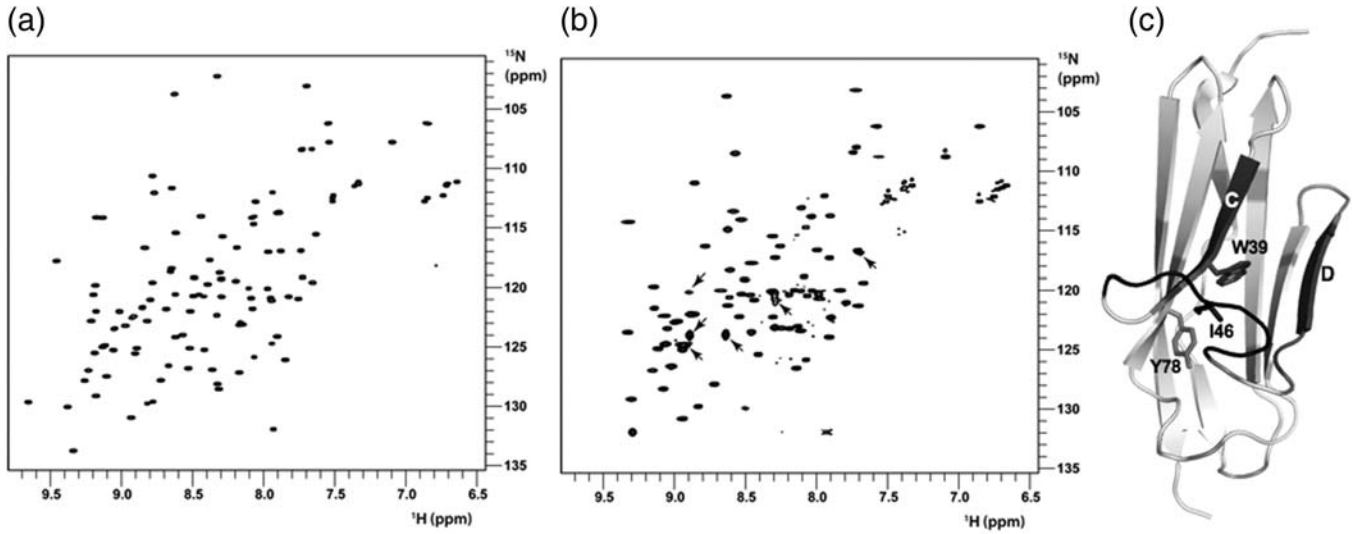

Fig. 3. Structural and dynamical comparison of Z1 and Z1^{FLAG}. [^{1}H,^{15}N]-HSQC spectra of (**a**) Z1 and (**b**) Z1^{FLAG}. Resonance assignments are shown in Supplementary Fig. S2. Low-intensity, broad, unassigned cross peaks in the Z1^{FLAG} spectrum are marked with arrows; (**c**) Resonance broadening in the Z1^{FLAG} [^{1}H,^{15}N]-HSQC spectrum mapped onto the structure of Z1 (relative intensities calculated as I_{Z1}^{FLAG}/I_{Z1}). The modified CD loop is in black. Residues unassigned due to severe exchange broadening are confined to β-strands C and D and shown in dark grey. A decrease in relative intensity by a factor >2.5 respect to the average value is observed only for sparse groups and is indicated in medium grey. (This figure is provided in color as Supplementary Fig. S5.)

crystallographic model; Marino *et al*., 2006; Zou *et al*., 2006). In the grafted loop, a decrease in line width signified an increase in internal dynamics, indicating that it had now disengaged from the β-structure and, in agreement with expectations, formed a structurally disordered loop.

The structural effect of the modification was further evaluated by comparison of ^{13}CO, $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ chemical shifts that report on the backbone dihedral angles. These values were very similar for most residues in Z1 and Z1^{FLAG}, with the average difference <0.3 ppm (Fig. 4a). Significant differences (>0.5 ppm) were few and vicinal to β-strands C and D (Fig. 4b). To translate the differences in chemical shifts into differences in structure, we predicted the values of the backbone dihedral angles for residues in Z1 and Z1^{FLAG} from their respective chemical shifts and calculated the difference. For most residues, the values were within 10° of each other, with even the largest difference below 20° (Fig. 4c and d), in agreement with the small differences in chemical shifts. Notably, the predicted differences in dihedral angles were significantly smaller than the precision of the prediction, which was 20–30° for most residues. This precision is typical of DANGLE (Cheung *et al*., 2010), which estimates the variation of dihedral angles from a fragment database. However, when comparing proteins with localized mutations and small changes in chemical shift values, such precision estimation is misleading, as it suggests relatively large variation in dihedral angles even if all the chemical shifts are identical. Thus, the direct comparison of the predicted values as a measure of dihedral angle variation is more appropriate for the analysis of the effect of the mutation in this study. For the grafted FLAG sequence, the angle prediction was unreliable, correlating with a lack of defined conformation and noticeable mobility. The small predicted variations in backbone dihedral angles demonstrated that the secondary structure and, thus, the fold of Z1 was largely unaffected by the CD loop modification. In summary, in agreement with deductions from FTIR data, NMR showed that the effect of grafting the FLAG tag in the CD loop was limited to a certain exchange in β-strands C and D, but that this did not disrupt the fold.

Functionality of exogenous peptides in Z1

The functionality of the peptides displayed in the CD loop was tested using pull-down assays. For Z1^{FLAG}, the purified sample was incubated with anti-FLAG M2 agarose beads and eluted with free FLAG peptide. The variant clearly interacted with the resin, while wild-type Z1 did not reveal any detectable retention (Fig. 5a). This prompted us to study further the efficiency of the binding with the anti-FLAG antibody M2 in solution using isothermal titration calorimetry (ITC) (Fig. 5b). This yielded an estimated K_d of 1.47 ± 21.5 μM and the expected stoichiometry of 1 : 1, showing that the bioreactivity of the inserted peptide is in a scale comparable to that of general biological processes.

In contrast, a pull-down assay that used Z1^{PxxP} versus the SH3 domain of Pexp13 failed to reveal binding. Size-exclusion chromatography further showed that the samples did no co-emigrate but segregated as individual species, confirming the absence of significant interaction. This result might be due to the substantial length of the recognition sequence and the conformational requirements of the binding in this case (Douangamath *et al*., 2002). In Z1, the end points of the CD loop are ∼20 Å apart, so that the inserted 14 aa peptide might be forced into a bended structure that might preclude recognition. In addition, the high content of hydrophobic residues in the peptide could have caused a certain collapse of the engineered loop onto itself or onto the β-sandwich, so that the sequence might not be available for interaction. The rational introduction of such complex binders into protein scaffolds is difficult and this goal might be better achieved through display/selection technologies.

In conclusion, intracellular Ig domains are robust protein folds that are undemanding to produce recombinantly in bacterial cultures. The titin myofilament is a rich source of such Ig frames and a number of domains (see above) are

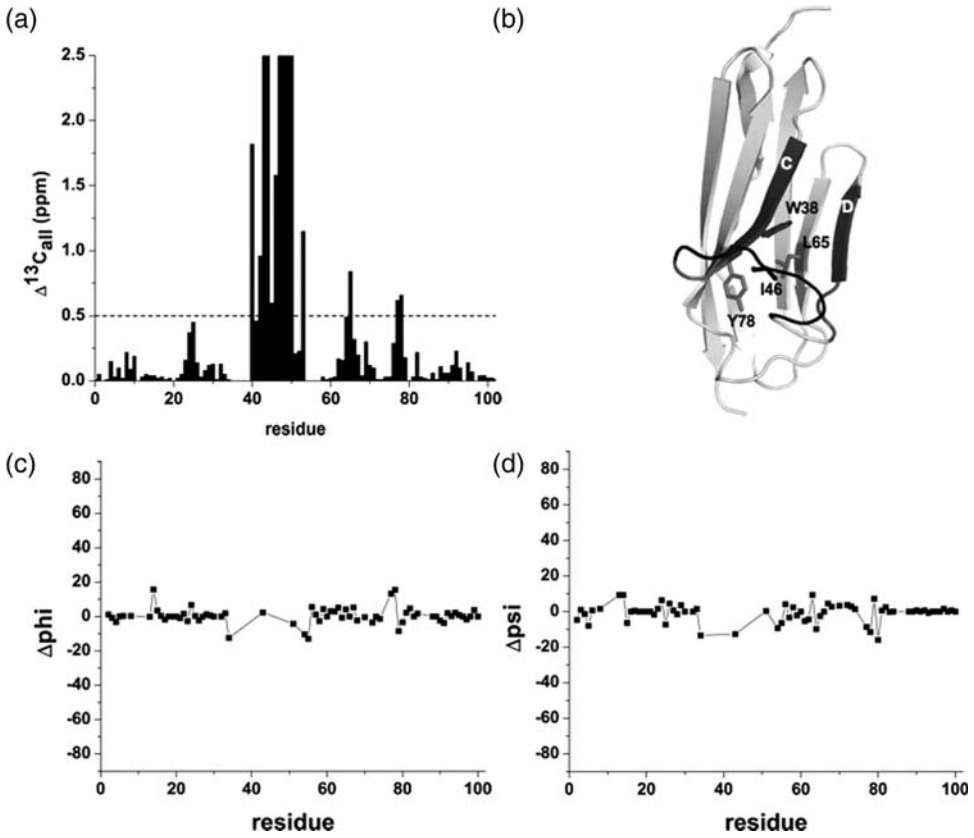

Fig. 4. Dihedral angle analysis from ^{13}C chemical values. (**a**) Average difference between the values of ^{13}CO, $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ chemical shifts of Z1 and $Z1^{FLAG}$; (**b**) Mapping of chemical shift differences onto the structure of Z1. The modified CD loop is in black. Residues unassigned in β-strands C and D due to broadening are in dark grey and average chemical shift differences >0.5 ppm are in medium grey; (**c** and **d**) differences in φ and ψ backbone dihedral angles (in degrees) predicted from the chemical shifts of Z1 and $Z1^{FLAG}$. (This figure is provided in color as Supplementary Fig. S6.)

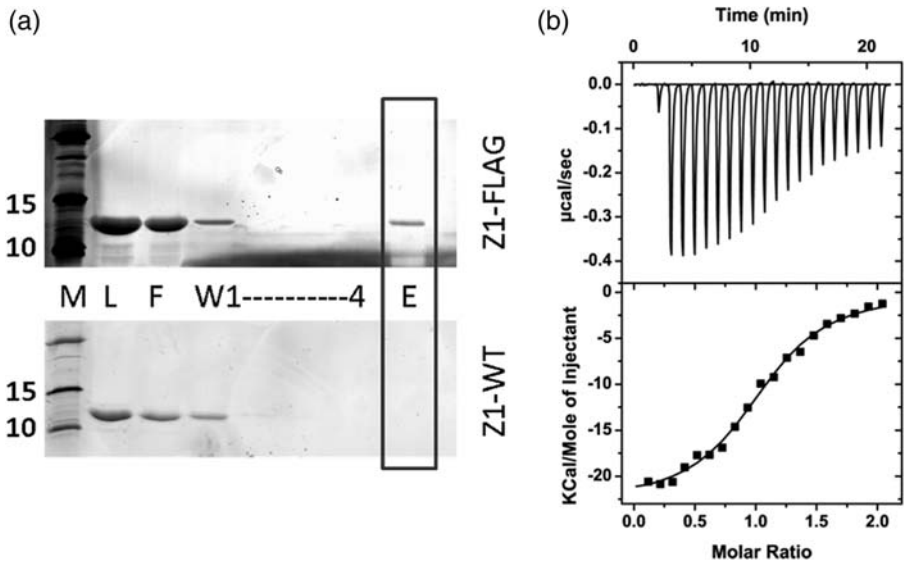

Fig. 5. Molecular binding properties of $Z1^{FLAG}$. (**a**) Pull-down assay using anti-FLAG M2 affinity gel. The SDS-PAGE shows molecular weight markers (M), incubated samples (L), unbound (F: flow-through, W1-4: washes) and bound fraction (E: elution); (**b**) ITC data on the interaction to anti-FLAG antibody.

now well-characterized that constitute a valuable diversity for experimentation. Intracellular I-Ig contain regions of natural variability and, in particular, the CD loop tolerates substantial modifications and is suitable for the engineering of recognition elements into the fold. It is possible that the adjacent, naturally variable EF loop might also tolerate diversification and that, together, both loops might permit developing complex 3D binding interfaces for molecular partners. Interestingly, the CD loop is remote from Ig–Ig interfaces in the natural poly-domain arrays of titin as well as from the self-assembling interface of a synthetic titin-based polymer recently reported (Bruning *et al.*, 2010).

Thus, it is tantalizing to envision the intracellular I-Ig fold as a promising protein scaffold for the development of protein polymers with controlled functions that might allow to bridge traditional Ig-based technologies and higher order assemblies in biotechnology.

Supplementary data

Supplementary data are available at *PEDS* online.

Funding

This work was supported by the Swiss National Foundation (3100A0-112595). B.F. was funded by a BBSRC studentship.

References

Barth,A. (2007) *Bioch. Biophys. Acta*, **1767**, 1073–1101.
Bruning,M., Leopoldseder,S., Müller,S.A., *et al*. (2010) *Nano Lett.*, **10**, 4533–4537.
Bucher,R.M., Svergun,D.I., Muhle-Goll,C. and Mayans,O. (2010) *J. Mol. Biol.*, **401**, 843–853.
Chami,M., Franke,B., Mayans,O. and Kreplak,L. (2011) *J. Biosc. Bioeng.*, **112**, 178–179.
Charles,P.T., Goldman,E.R., Rangasammy,J.G., Schauer,C.L., Chen,M.S. and Taitt,C.R. (2004) *Biosens Bioelectron.*, **20**, 753–764.
Cheung,M.S., Maguire,M.L., Stevens,T.J. and Broadhurst,R.W. (2010) *J Magn Reson.*, **202**, 223–233.
Chi,Y., Kumar,T.K., Wang,H.M., Ho,M.C., Chiu,I.M. and Yu,C. (2001) *Biochemistry*, **40**, 7746–7753.
Deprez,C., Lloubes,R., Gavioli,M., Marion,D., Guerlesquin,F. and Blanchard,L. (2005) *J. Mol. Biol.*, **346**, 1047–1057.
Douangamath,A., Filipp,F.V., Klein,A., *et al*. (2002) *Mol. Cell*, **10**, 1007–1017.
Edlich,C. and Muhle-Goll,C. (2003) *J. Biomol. NMR*, **27**, 283–284.
Gebauer,M. and Skerra,A. (2009) *Curr. Opin. Chem. Biol.*, **13**, 245–255.
Goll,C.M., Pastore,A. and Nilges,M. (1998) *Structure*, **6**, 1291–1302.
Grönwall,C. and Ståhl,S. (2009) *J. Biotechnol.*, **140**, 254–269.
Halaby,D.M., Poupon,A. and Mornon,J. (1999) *Protein Eng.*, **12**, 563–571.
Hamill,S.J., Cota,E., Chothia,C. and Clarke,J. (2000) *J. Mol. Biol.*, **295**, 641–649.
Holliger,P. and Hudson,P.J. (2005) *Nat. Biotechnol.*, **23**, 1126–1136.
Improta,S., Politou,A.S. and Pastore,A. (1996) *Structure*, **4**, 323–337.
Ioerger,T.R., Du,C. and Linthicum,D.S. (1999) *Mol. Immunol.*, **36**, 373–386.
Kontrogianni-Konstantopoulos,A., Ackermann,M.A., Bowman,A.L., Yap,S.V. and Bloch,R.J. (2009) *Physiol. Rev.*, **89**, 1217–1267.
Krissinel,E. and Henrick,K. (2004) *Acta Crystallogr.*, **D60**, 2256–2268.
Labeit,S. and Kolmerer,B. (1995) *Science*, **270**, 293–296.
Lo,A.S., Zhu,Q. and Marasco,W.A. (2008) *Handb. Exp. Pharmacol.*, **181**, 343–373.
Löfblom,J., Frejd,F.Y. and Ståhl,S. (2011) *Curr. Opin. Biotechnol.*, **22**, 843–848.
Marino,M., Svergun,D.I., Kreplak,L., Konarev,P.V., Maco,B., Labeit,D. and Mayans,O. (2005) *J. Muscle Res. Cell Mot.*, **26**, 355–365.
Marino,M., Zou,P., Svergun,D.I., Garcia,P., Edlich,C., Simon,B., Wilmanns,M., Muhle-Goll,C. and Mayans,O. (2006) *Structure*, **14**, 1437–1447.
Mayans,O., Wuerges,J., Canela,S., Gautel,M. and Wilmanns,M. (2001) *Structure*, **9**, 331–340.
Mrosek,M., Labeit,D., Witt,S., Heerklotz,H., von Castelmur,E., Labeit,S. and Mayans,O. (2007) *FASEB J.*, **21**, 1383–1392.
Müller,S., Lange,S., Gautel,M. and Wilmanns,M. (2007) *J. Mol. Biol.*, **371**, 469–480.
Nelson,A.L. and Reichert,J.M. (2009) *Nat. Biotech.*, **27**, 331–337.
Niesen,F.H., Berglund,H. and Vedadi,M. (2007) *Nat. Protoc.*, **2**, 2212–2221.
Pei,J., Kim,B.H. and Grishin,N.V. (2008) *Nucleic Acids Res.*, **36**, 2295–2300.
Pernigo,S., Fukuzawa,A., Bertz,M., Holt,M., Rief,M., Steiner,R.A. and Gautel,M. (2010) *Proc. Natl Acad. Sci. USA*, **107**, 2908–2913.
Pfuhl,M. and Pastore,A. (1995) *Structure*, **3**, 391–401.
Politou,A.S., Thomas,D.J. and Pastore,A. (1995) *Biophys. J.*, **69**, 2601–2610.
Steinmeyer,D.E. and McCormick,E.L. (2008) *Drug Discov. Today*, **13**, 613–618.
von Castelmur,E., Marino,M., Svergun,D.I., *et al*. (2008) *Proc. Natl Acad. Sci. USA*, **105**, 1186–1191.
Vranken,W.F., Boucher,W., Stevens,T.J., *et al*. (2005) *Proteins*, **59**, 687–696.
Wishart,D.S., Bigam,C.G., Yao,J., Abildgaard,F., Dyson,H.J., Oldfield,E., Markley,J.L. and Sykes,B.D. (1995) *J. Biomol. NMR*, **6**, 135–140.
Wright,C.F., Christodoulou,J., Dobson,C.M. and Clarke,J. (2004) *Protein Eng. Des. Sel.*, **17**, 443–453.
Wurch,T., Lowe,P., Caussanel,V., Bes,C., Beck,A. and Corvaia,N. (2008) *Curr. Pharm. Biotechnol.*, **9**, 502–509.
Zou,P., Pinotsis,N., Lange,S., Song,Y.H., Popov,A., Mavridis,I., Mayans,O., Gautel,M. and Wilmanns,M. (2006) *Nature*, **439**, 229–233.